

4 6

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/900,751

DATE: 10/29/2001

TIME: 17:58:04

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF3\10292001\I900751.raw

```

4 <110> APPLICANT: Allen, Keith D
5     Leviten, Michael W.
7 <120> TITLE OF INVENTION: TRANSGENIC MICE CONTAINING SERINE
8     PROTEASE GENE DISRUPTIONS
10 <130> FILE REFERENCE: R-386
12 <140> CURRENT APPLICATION NUMBER: US 09/900,751
13 <141> CURRENT FILING DATE: 2001-07-06
15 <150> PRIOR APPLICATION NUMBER: US 60/217,449
16 <151> PRIOR FILING DATE: 2000-07-10
18 <150> PRIOR APPLICATION NUMBER: US 60/223,170
19 <151> PRIOR FILING DATE: 2000-08-07
21 <150> PRIOR APPLICATION NUMBER: US 60/223,460
22 <151> PRIOR FILING DATE: 2000-08-07
24 <160> NUMBER OF SEQ ID NOS: 4
26 <170> SOFTWARE: FastSEQ for Windows Version 4.0
28 <210> SEQ ID NO: 1
29 <211> LENGTH: 3106
30 <212> TYPE: DNA
31 <213> ORGANISM: Mus musculus
33 <400> SEQUENCE: 1
34 catggtagac ggctgcccgg agggaccacg cgtctgagac cggcgatcgg accgccaaaa 60
35 ccatgggtag caatcggggc cgcaaggccg gagggggctc tcaggacttc ggcgcggggac 120
36 tcaagtacaa ctcccggcta gagaacatga atggctttga ggaggggtgt gagttcctgc 180
37 ctgcgaacaa tgccaagaaa gtggagaagc gagggcccag gcgctgggtg gtgctgggtg 240
38 cagtgtgtgt cagcttcttc ttgctctccc tcatggctgg cttgctgggtg tggcacttcc 300
39 attatcgga tgtgcgggtt caaaaagtct tcaatggcca tctgaggatc acaaatgaga 360
40 tctttctgga tgcgtatgag aactccacct ccacagagtt tatcagcctg gccagccagg 420
41 tgaaggaggc gctgaagctg ctgtacaatg aagtccctgt cctgggtccc taccacaaga 480
42 agtcggctgt aactgccttc agtgaggcca gtgtcatcgc ctactactgg tcagagttca 540
43 gcatccccc acacctggca gaagagggtg atcgcgccat ggctgtggag cgagttgtaa 600
44 cattgccacc ccgagcacgg gcaactgaaat ccttcgtgct aacatctgtg gtggccttcc 660
45 ccattgaccc cagaatgctg cagaggactc aggacaacag ctgcagtttt gccctgcatg 720
46 cccatggtgc agcagtga caacttcaacta cccctggctt ccccaacagt ccttaccgg 780
47 cgcagtcggc ctgcccagtg gtcttgccgg gggacgcoga ctctgtgctg agcctcacct 840
48 tccgaagctt tgatgtcgct cctgtgatg agcatggcag tgacctggc accgtgtatg 900
49 atagcctgag ccccatggaa ccccaagctg tgggtgcggc gtgtggcacc ttctcacct 960
50 cctacaacct gactttcttc tctctccaga acgtcttctt tgtcacgctg ataaccaata 1020
51 ctgaccggcg acatcctggc tttgaggcca ctttcttcca gctgcccagg atgagcagct 1080
52 gtggcggctt tttgagtga acccaagggg catttagcag cccctactat ccaggccact 1140
53 acccgcccaa catcaactgc acatggaata tcaaggtgcc caacaaccgg aacgtgaagg 1200
54 tgcgttcaa actcttctat ctggtggacc ccaacgtacc agtgggctcc tgcaccaagg 1260
55 actatgtgga gatcaacggg gagaagtact gcggtgagag gtcccagttt gtggtgagca 1320
56 gcaacagcag caagattaca gtccacttcc attctgatca ctgtacacg gacaccgggt 1380
57 tcttagctga gtacctctcc tacgactcca acgaccgtg cccagggatg ttcattgtga 1440
58 agactggacg gtgcatccga aaggaactgc gctgcgacgg ctgggcagac tgcccggatt 1500
59 atagtgtatg gcgttactgc cgatgcaatg ccaccacca gtacacgtgc aaaaaccagt 1560
60 tctgcaagcc cctcttctgg gtctgtgaca gtgtcaacga ctgtggggac ggaagtgcag 1620

```

RAW SEQUENCE LISTING

DATE: 10/29/2001

PATENT APPLICATION: US/09/900,751

TIME: 17:58:04

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF3\10292001\I900751.raw

```

61 aggagggctg cagctgtcct gctgggagtt tcaagtgttc caatgggaag tgtctccctc 1680
62 agagccagaa gtgtaatggg aaggacaact gtggagatgg gtctgacgag gcttcatgtg 1740
63 acagcgtgaa tgtcgtctct tgcaccaa atacctaccg ctgccaaaat ggcctctgtc 1800
64 tgagcaaggg caacctgag tgtgatggga agacggactg tagcgatggc tccgatgaga 1860
65 aaaactgtga ctgtgggctg cgatccttta ccaaacaggc tcgctgggtt ggtggcacga 1920
66 atgcggacga gggcgagtgg ccctggcagg tgagcctcca cgccctgggc cagggccact 1980
67 tgtgtggggc ctgcgtcatc tctcctgact ggctgggtctc tgcagctcat tgctttcagg 2040
68 atgacaaaaa tttcaagtac tcagactaca cgatgtggac ggccttcctg ggtctgctgg 2100
69 accagagcaa gcgcagtgcc tctggggtgc aggagctgaa gctcaaactg atcatcacc 2160
70 acccttcctt caatgatatt accttcgact atgacatcgc ctgtctggag ctggagaagt 2220
71 cggaggagta cagcaccgtc gtgcgcccc a tctgcctgcc tgatgctacc catgtcttcc 2280
72 ctgctggcaa ggccatctgg gtcacaggct gggggcacac aaaagaggga ggtaccggag 2340
73 cgctgacctt gcagaagggt gagatccgtg tcatcaacca gaccacctgt gaggacctca 2400
74 tgccgcagca gatcacccca cgaatgatgt gtgtgggttt cctcagtggg ggtgtggact 2460
75 cctgccaggg tgactctggt ggcctcttgc caagcgagg gaaagatggg cgaatgttcc 2520
76 aggtctggtg ttgtagctgg ggtgaaggct gcgctcagag gaacaagcca ggcgtgtaca 2580
77 caaggctccc tgtagtctcg gactggatca aagagcacac tggggtatag cagcatggac 2640
78 agacagccga ccacaaacac ccacagggat gcccgcacatg cacacctgga tacaggagag 2700
79 gaacactgac gacatttatg ctgtggcctc cccccccaa cacaaccag actgtgaact 2760
80 gcacccctag gactcagagt tcttccaaag tgggaccccc tcaagagttg gagagagaac 2820
81 ttgcgtgcta gcggcccgac ctgggggcaa ggggttgatg gcagccttcc ccctctagcc 2880
82 ctgagctggg tgaagatgat gctgtcccg agagctgctt ccactgttca ttgagctccc 2940
83 gggagcccta tgggaggagg ggctcagggt cactcttttc aggaagcgcc agccctagga 3000
84 accccagaaa agagtggtag ctaaggctga aattgttttg ctgttgccag ggggtgggat 3060
85 ttgagagtaa aacattttat ttctttttta aaaaaaaaaa aaaaaa 3106

```

87 <210> SEQ ID NO: 2

88 <211> LENGTH: 855

89 <212> TYPE: PRT

90 <213> ORGANISM: Mus musculus

92 <400> SEQUENCE: 2

```

93 Met Gly Ser Asn Arg Gly Arg Lys Ala Gly Gly Gly Ser Gln Asp Phe
94 1 5 10 15
95 Gly Ala Gly Leu Lys Tyr Asn Ser Arg Leu Glu Asn Met Asn Gly Phe
96 20 25 30
97 Glu Glu Gly Val Glu Phe Leu Pro Ala Asn Asn Ala Lys Lys Val Glu
98 35 40 45
99 Lys Arg Gly Pro Arg Arg Trp Val Val Leu Val Ala Val Leu Phe Ser
100 50 55 60
101 Phe Leu Leu Leu Ser Leu Met Ala Gly Leu Leu Val Trp His Phe His
102 65 70 75 80
103 Tyr Arg Asn Val Arg Val Gln Lys Val Phe Asn Gly His Leu Arg Ile
104 85 90 95
105 Thr Asn Glu Ile Phe Leu Asp Ala Tyr Glu Asn Ser Thr Ser Thr Glu
106 100 105 110
107 Phe Ile Ser Leu Ala Ser Gln Val Lys Glu Ala Leu Lys Leu Leu Tyr
108 115 120 125
109 Asn Glu Val Pro Val Leu Gly Pro Tyr His Lys Lys Ser Ala Val Thr
110 130 135 140
111 Ala Phe Ser Glu Gly Ser Val Ile Ala Tyr Tyr Trp Ser Glu Phe Ser

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/900,751

DATE: 10/29/2001

TIME: 17:58:04

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF3\10292001\I900751.raw

```

112 145                      150                      155                      160
113 Ile Pro Pro His Leu Ala Glu Glu Val Asp Arg Ala Met Ala Val Glu
114                      165                      170                      175
115 Arg Val Val Thr Leu Pro Pro Arg Ala Arg Ala Leu Lys Ser Phe Val
116                      180                      185                      190
117 Leu Thr Ser Val Val Ala Phe Pro Ile Asp Pro Arg Met Leu Gln Arg
118                      195                      200                      205
119 Thr Gln Asp Asn Ser Cys Ser Phe Ala Leu His Ala His Gly Ala Ala
120                      210                      215                      220
121 Val Thr Arg Phe Thr Thr Pro Gly Phe Pro Asn Ser Pro Tyr Pro Ala
122 225                      230                      235                      240
123 His Ala Arg Cys Gln Trp Val Leu Arg Gly Asp Ala Asp Ser Val Leu
124                      245                      250                      255
125 Ser Leu Thr Phe Arg Ser Phe Asp Val Ala Pro Cys Asp Glu His Gly
126                      260                      265                      270
127 Ser Asp Leu Val Thr Val Tyr Asp Ser Leu Ser Pro Met Glu Pro His
128                      275                      280                      285
129 Ala Val Val Arg Leu Cys Gly Thr Phe Ser Pro Ser Tyr Asn Leu Thr
130                      290                      295                      300
131 Phe Leu Ser Ser Gln Asn Val Phe Leu Val Thr Leu Ile Thr Asn Thr
132 305                      310                      315                      320
133 Asp Arg Arg His Pro Gly Phe Glu Ala Thr Phe Phe Gln Leu Pro Lys
134                      325                      330                      335
135 Met Ser Ser Cys Gly Gly Phe Leu Ser Asp Thr Gln Gly Thr Phe Ser
136                      340                      345                      350
137 Ser Pro Tyr Tyr Pro Gly His Tyr Pro Pro Asn Ile Asn Cys Thr Trp
138                      355                      360                      365
139 Asn Ile Lys Val Pro Asn Asn Arg Asn Val Lys Val Arg Phe Lys Leu
140                      370                      375                      380
141 Phe Tyr Leu Val Asp Pro Asn Val Pro Val Gly Ser Cys Thr Lys Asp
142 385                      390                      395                      400
143 Tyr Val Glu Ile Asn Gly Glu Lys Tyr Cys Gly Glu Arg Ser Gln Phe
144                      405                      410                      415
145 Val Val Ser Ser Asn Ser Ser Lys Ile Thr Val His Phe His Ser Asp
146                      420                      425                      430
147 His Ser Tyr Thr Asp Thr Gly Phe Leu Ala Glu Tyr Leu Ser Tyr Asp
148                      435                      440                      445
149 Ser Asn Asp Pro Cys Pro Gly Met Phe Met Cys Lys Thr Gly Arg Cys
150                      450                      455                      460
151 Ile Arg Lys Glu Leu Arg Cys Asp Gly Trp Ala Asp Cys Pro Asp Tyr
152 465                      470                      475                      480
153 Ser Asp Glu Arg Tyr Cys Arg Cys Asn Ala Thr His Gln Phe Thr Cys
154                      485                      490                      495
155 Lys Asn Gln Phe Cys Lys Pro Leu Phe Trp Val Cys Asp Ser Val Asn
156                      500                      505                      510
157 Asp Cys Gly Asp Gly Ser Asp Glu Glu Gly Cys Ser Cys Pro Ala Gly
158                      515                      520                      525
159 Ser Phe Lys Cys Ser Asn Gly Lys Cys Leu Pro Gln Ser Gln Lys Cys
160                      530                      535                      540

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/900,751

DATE: 10/29/2001

TIME: 17:58:04

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF3\10292001\I900751.raw

```

161 Asn Gly Lys Asp Asn Cys Gly Asp Gly Ser Asp Glu Ala Ser Cys Asp
162 545 550 555 560
163 Ser Val Asn Val Val Ser Cys Thr Lys Tyr Thr Tyr Arg Cys Gln Asn
164 565 570 575
165 Gly Leu Cys Leu Ser Lys Gly Asn Pro Glu Cys Asp Gly Lys Thr Asp
166 580 585 590
167 Cys Ser Asp Gly Ser Asp Glu Lys Asn Cys Asp Cys Gly Leu Arg Ser
168 595 600 605
169 Phe Thr Lys Gln Ala Arg Val Val Gly Gly Thr Asn Ala Asp Glu Gly
170 610 615 620
171 Glu Trp Pro Trp Gln Val Ser Leu His Ala Leu Gly Gln Gly His Leu
172 625 630 635 640
173 Cys Gly Ala Ser Leu Ile Ser Pro Asp Trp Leu Val Ser Ala Ala His
174 645 650 655
175 Cys Phe Gln Asp Asp Lys Asn Phe Lys Tyr Ser Asp Tyr Thr Met Trp
176 660 665 670
177 Thr Ala Phe Leu Gly Leu Leu Asp Gln Ser Lys Arg Ser Ala Ser Gly
178 675 680 685
179 Val Gln Glu Leu Lys Leu Lys Arg Ile Ile Thr His Pro Ser Phe Asn
180 690 695 700
181 Asp Phe Thr Phe Asp Tyr Asp Ile Ala Leu Leu Glu Leu Glu Lys Ser
182 705 710 715 720
183 Val Glu Tyr Ser Thr Val Val Arg Pro Ile Cys Leu Pro Asp Ala Thr
184 725 730 735
185 His Val Phe Pro Ala Gly Lys Ala Ile Trp Val Thr Gly Trp Gly His
186 740 745 750
187 Thr Lys Glu Gly Gly Thr Gly Ala Leu Ile Leu Gln Lys Gly Glu Ile
188 755 760 765
189 Arg Val Ile Asn Gln Thr Thr Cys Glu Asp Leu Met Pro Gln Gln Ile
190 770 775 780
191 Thr Pro Arg Met Met Cys Val Gly Phe Leu Ser Gly Gly Val Asp Ser
192 785 790 795 800
193 Cys Gln Gly Asp Ser Gly Gly Pro Leu Ser Ser Ala Glu Lys Asp Gly
194 805 810 815
195 Arg Met Phe Gln Ala Gly Val Val Ser Trp Gly Glu Gly Cys Ala Gln
196 820 825 830
197 Arg Asn Lys Pro Gly Val Tyr Thr Arg Leu Pro Val Val Arg Asp Trp
198 835 840 845
199 Ile Lys Glu His Thr Gly Val
200 850 855
203 <210> SEQ ID NO: 3
204 <211> LENGTH: 200
205 <212> TYPE: DNA
206 <213> ORGANISM: Artificial Sequence
208 <220> FEATURE:
209 <223> OTHER INFORMATION: Targeting vector
211 <400> SEQUENCE: 3
212 ttccccattg agactggctt accccggaag ctgcctgcct cagtctcccg cttcctgtct 60
213 cccaggtac cggagcgctg atcctgcaga agggtagat ccgtgtcatc aaccagacca 120

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/900,751

DATE: 10/29/2001

TIME: 17:58:04

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF3\10292001\I900751.raw

```
214 cctgtgagga cctcatgccg cagcagatca cccacgaat gatgtgtgtg ggtttcctca 180
215 gtgggggtgt ggactcctgc 200
217 <210> SEQ ID NO: 4
218 <211> LENGTH: 200
219 <212> TYPE: DNA
220 <213> ORGANISM: Artificial Sequence
222 <220> FEATURE:
223 <223> OTHER INFORMATION: Targeting vector
225 <400> SEQUENCE: 4
226 atgggcgaat gttccaggct ggtgtggtgg gctgggggtga aggctgcgct cagaggaaca 60
227 agccaggcgt gtacacaagg ctccctgtag ttcgggactg gatcaaagag cacactgggg 120
228 tatagcagca tggacagaca gccgaccaca aacaccaca gggatgcccg acatgcacac 180
229 ctggatacag gagagggaca 200
```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/900,751

DATE: 10/29/2001

TIME: 17:58:05

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF3\10292001\I900751.raw